

⑫ 公開特許公報(A)

昭61-97790

⑮ Int.Cl.⁴

G 06 K 9/62

識別記号

庁内整理番号

Z-8320-5B

⑬ 公開 昭和61年(1986)5月16日

審査請求 未請求 発明の数 2 (全4頁)

⑭ 発明の名称 オンライン文字認識装置

⑯ 特 願 昭59-219030

⑰ 出 願 昭59(1984)10月18日

⑱ 発 明 者 亘 理 誠 夫 東京都港区芝5丁目33番1号 日本電気株式会社内

⑲ 出 願 人 日本電気株式会社 東京都港区芝5丁目33番1号

⑳ 代 理 人 弁理士 内 原 晋

明 細 書

1. 発明の名称

オンライン文字認識装置

2. 特許請求の範囲

(1) 文字の時系列データを発生する文字データ発生部と、文字を書く際の、あるストロークを書き終えてから次のストロークを書き始めるまでの時間を計測する計時部と、前記計時部において計測された時間を読み込みタイムアウト値を逐次学習するタイムアウト学習部と、前記計時部において計測された時間が前記タイムアウト学習部で求められたタイムアウト値を超えたとき一文字の書き終りと判断する判定部と、前記文字データ発生部から出力される文字データより文字を認識する文字認識部とを有することを特徴とするオンライン文字認識装置。

(2) 文字の時系列データを発生する文字データ発生部と、文字を書く際の、あるストロークを書き終ってから次のストロークを書き始めるまでの時間を計測する計時部と、文字の字種を選択する

字種選択部と、前記計時部において計測された時間を読み込み前記字種選択部より指定された字種ごとにタイムアウト値を逐次学習するタイムアウト学習部と、前記計時部において計測された時間が前記タイムアウト学習部で求められたタイムアウト値を超えたとき一文字の書き終りと判断する判定部と、前記文字データ発生部から出力される文字データより文字を認識する文字認識部とを有することを特徴とするオンライン文字認識装置。

3. 発明の詳細な説明

(産業上の利用分野)

本発明は、実時間で入力された筆跡をもとに文字を認識するオンライン文字認識装置の改良に関するものである。

(従来技術)

従来、実時間で入力された筆跡をもとに文字を認識する手法は既に存在する。例えば特願昭56-19417号明細書(以下文献(1)と称す)に述べられているようなボタンマッチングを用いた手法が有効である。文字を認識する場合、書かれた筆跡

の内どこからどこまでが一つの文字であるかを決定しなければならない。この決定方法には従来以下に述べる方法が知られていた。

ストロークとストロークの時間、すなわちペン先がタブレット上から離れてから次にタブレット上に着くまでの時間を計測して、この時間がある定められた閾値を越えた時、タイムアウトとしてこの文字の書き終りとする方法である。

(発明が解決しようとする問題点)

しかし、前記方法ではタイムアウト値は書く人の特性例えばゆっくりと文字が書かれる場合や、文字の途中でペンを休めることがある人といった条件を考慮した値に設定しなければならない。なぜならタイムアウトを小さく設定すると書いてる途中でペンを休めたときに、そのところで文字の終りと判断され、文字を正しく入力できないからである。一方、タイムアウトを大きく設定すると文字を書き終ってから、認識結果が表示されるまで待たなければならないという欠点になる。

本発明の目的は、過去に書かれた文字のストロ

クを書き始めるまでの時間を計測する計時部2と、文字の字種を選択する字種選択部6と、前記計時部2において計測された時間を読み込み前記字種選択部6より指定された字種ごとにタイムアウト値を逐次学習するタイムアウト学習部3と、前記計時部2において計測された時間が前記タイムアウト学習部3で求められたタイムアウト値を越えたとき一文字の書き終りと判断する判定部4と、前記文字データ発生部から出力される文字データより文字を認識する文字認識部5とを有するオンライン文字認識装置である。

(発明の原理・作用)

次に本発明の原理を説明する。オンライン文字認識装置ではストローク間の時間がタイムアウト値を越えた時、その文字の書き終りとしており、タイムアウト値は一文字が分割されない範囲で小さい値に設定するのが望ましい。このため、本発明では一般的なタイムアウト値を初期値として設定し、一文字づつ認識すると共に、過去n文字のストローク間の時間 t_{i-n} によりタイムアウト

ーク間の時間を用いてタイムアウト値を逐次学習し変更することによって、書く人に適応したタイムアウト値を設定し文字の書き終りの検出精度を上げて、オンライン文字認識装置の使い勝手を改善することにある。

(問題点を解決するための手段)

本発明は、文字の時系列データを発生する文字データ発生部1と、文字を書く際のあるストロークを書き終ってから次のストロークを書き始めるまでの時間を計測する計時部2と、前記計時部2において計測された時間を読み込みタイムアウト値を逐次学習するタイムアウト学習部3と、前記計時部2において計測された時間が前記タイムアウト学習部3で求められたタイムアウト値を越えたとき一文字の書き終りと判断する判定部4と、前記文字データ発生部1から出力される文字データより文字を認識する文字認識部5とを有するオンライン文字認識装置および文字の時系列データを発生する文字データ発生部1と、文字を書く際のあるストロークを書き終えてから次のストロ

値を修正する。すなわち

初期値 $T_0 = T_s$ とし

i番目のタイムアウト値は

$$T_i = T_{i-1} + f(t_{i-1}, t_{i-2}, \dots, t_{i-n}) \dots (1)$$

で与えられる。ここで学習関数fは

$$f(t_{i-1}, t_{i-2}, \dots, t_{i-n}) =$$

$$\frac{1}{n} \sum_{j=1}^n (t_{i-j} - T_{i-1}) \dots (2)$$

を用いることが考えられる。これは過去n文字のストローク間の時間とタイムアウト値との平均誤差のn倍を加算することによりタイムアウト値を修正している。

以上説明したようにタイムアウト値は逐次学習により自動的に修正され、利用者に適応したタイムアウト値が設定できる。

しかし、本発明の特許請求の範囲第1項の発明は字種の選択をせず、例えば漢字もカタカナも同一のタイムアウト値が設定したものであるが、特許請求の範囲第2項の発明では、字種選択部を設け利用者による字種選択情報を用いてタイムアウト値を字種ごとに学習するようになっている。

すなわち、タイムアウトの初期値は字種ごとに複数個あり、(1)式の計算は字種ごとに求められる。以上説明したようにタイムアウト値は字種ごとに自動的に修正され、字種に適応したタイムアウト値を設定できるものである。

(実施例1)

次に本発明の第1の実施例を図面を用いて説明する。本実施例は第1図に示すように構成される。文字データ発生部1より発生した文字の筆跡情報Cは、文字認識部5に入力されて保持される。文字データ発生部1としてはタブレットを使用することができる。このタブレットによる文字データ発生部1に得られたペンのアップ、ダウン情報Pが計時部2に入力される。計時部2では、この情報Pをもとにペンがアップになってから次にダウンになるまでの時間、すなわちストローク間の時間 Δt_i が計測される。タイムアウト学習部3では、ストローク間の時間 Δt_i をもとに前記(1)式を計算し、タイムアウト値 T_i を求める。ここに初期値 T_s 、係数 α は定数であらかじめ設定された値である。

場合は画数が少く途中でペンを休めることはほとんどないためタイムアウト値は小さい。このためタイムアウト学習部3では字種ごとにタイムアウト値を学習する。すなわち字種の数だけタイムアウト初期値 T_s を持ち字種ごとに(1)式に従ってタイムアウト値 T_i を求める。判定部4ではこのタイムアウト値 T_i とストローク間の時間 Δt_i と比較し認識結果出力可信号Rを出力する。文字認識部5では、選択された字種のみを認識対象とし筆跡情報Cをもとに文字が認識される。

以上、本発明による実施例に関して述べたが本発明はこの実施例にのみ限定されるものではない。

タイムアウト値の学習は(1)、(2)により求めているが、(2)式の代りに(3)式の過去 n 個の最大値を用いる方法

$$f(\Delta t_{i-1}, \Delta t_{i-2}, \dots, \Delta t_{i-n}) = \max(\Delta t_{i-1}, \Delta t_{i-2}, \dots, \Delta t_{i-n}) \dots (3)$$

また(1)式の代りに(4)式の逐次修正でなく置換してしまう方法

$$T_{i-1} = f(\Delta t_{i-1}, \Delta t_{i-2}, \dots, \Delta t_{i-n}) \dots (4)$$

判定部4では、タイムアウト値 T_i とストローク間の時間 Δt_i を比較し、 $\Delta t_i > T_i$ となった時認識結果出力可信号Rが文字認識部5へ出力される。文字認識部5では、ペンがアップした時刻よりそれまでに入力された筆跡情報Cをもとに文字が認識される。そして認識結果出力可信号Rが出力された時刻より後に認識結果が出力される。このような文字認識部2には、前記文献(1)に述べられているようなパターンマッチングを用いた手法を使用する。

(実施例2)

次に本発明第2の実施例を説明する。

第2の実施例は第2図に示すように構成され、第1の実施例に字種選択部6を付加したものである。字種選択部6では、利用者が字種を選択し、選択された字種Sがタイムアウト学習部3と文字認識部5とへ伝えられる。字種としては例えば漢字、カタカナ2種を考えるものとする。漢字の場合は通常画数が多く途中でペンを休めることもあるためタイムアウト値は大きく、一方カタカナの

などを用いることができる。

さらに、第2の実施例で字種は2種を用いているが、もっと細かく分けることもできる。

本発明は文字枠により文字の書き終りを検出する方法と共に用いることもできる。すなわち、タブレット上にいくつかの文字枠を設けておき、書かれていた文字枠以外に書き始めた場合前の文字が完結したとする方法である。この時まだ次の文字枠に書かれていなくてもストローク間の時間がタイムアウト値を越えた場合は文字の書き終りとする。

さらに、文字認識部5における文字認識方式は実施例で用いたパターンマッチング方式に限定されるものではない。

(発明の効果)

以上のように本発明によれば、過去に書かれた文字のストローク間の時間を用いて書く人に適応したタイムアウト値を設定することができ、一文字の途中でタイムアウトとなることもなく、かつ文字を書き終った後結果も速く表示することがで

き、オンライン文字認識装置の使い勝手を向上できる。

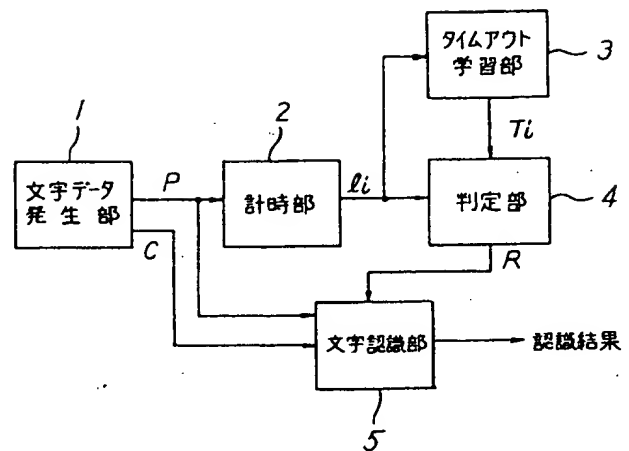
また、本発明によれば、字種選択情報を利用して字種ごとにタイムアウト値を学習することができ、利用者の書く字種ごとの速さに適応してタイムアウトを設定することにより、漢字を書いている途中で少し休んでもタイムアウトとはならず、またカタカナを書いている時に結果の表示に待たされることもなくなり、オンライン文字認識装置の使い勝手をより向上できる効果を有するものである。

4. 図面の簡単な説明

第1図は本発明の第1の実施例を示すブロック図、第2図は本発明の第2の実施例を示すブロック図である。

- 1 …… 文字データ発生部、 2 …… 計時部、
 3 …… タイムアウト学習部、 4 …… 判定部、
 5 …… 文字認識部、 6 …… 字種選択部

第1図



第2図

